https://doi.org/10.65025/ICAIC25128s

Machine learning for financial statement analysis and forecast: A case study of Guaranty Trust Holding Company PLC (GTCO PLC.).

Ayomide Sobowale
Final University, Cyprus
atemmy01@gmail.com

Abstract
This thesis explores the integration of machine learning techniques into financial statement analysis and forecasting within the context of the Nigerian Banking Industry. The primary objectives are to predict the financial health and profitability of companies, enhance data-driven investment decisions, reduce the risk of financial losses for investors, and advance the application of machine learning in finance by substituting traditional analysis methods. The study focuses on GTCO Plc, a prominent entity listed on the Nigerian Stock Exchange. Initially, traditional annual report analysis serves as a foundation for understanding the company's financial performance. Subsequently, a predictive model is developed using four machine learning algorithms: Random Forest, K-Nearest Neighbor, Logistics Regression, and Naïve Bayes. The outcomes of this research contribute to a nuanced understanding of machine learning's efficacy in financial analysis, offering a potential paradigm shift from conventional methodologies. The proposed model not only aids in predicting financial health and profitability but also empowers investors with valuable insights, mitigating financial risks in the dynamic Nigerian Banking Industry. This study thus marks a pivotal step towards fostering a data-driven investment landscape and embracing machine learning applications in financial decision-making.

Keywords
Financial Statement, Annual Report, Machine Learning Algorithms, Machine Learning Models, Random Forest, K-Nearest Neighbor, Logistics Regression, Naïve Bayes

1. Introduction

Every individual has a quest for financial freedom, and our desire of having several investment portfolios is increasing due to the longing for financial liberty. We often admire investors like Warren Buffett, Benjamin Graham, who have become successful in investment with a proper understanding of the market and by even setting investment rules to guide his decisions. Sadly, many investors today do not have adequate financial acumen thereby leading them to making bad decision which leads to loses. In order to limit investment loses, investors should fully understand and be able to analyse a company's annual report.

The analysis of financial statements is used to evaluate the financial performance of an organisation. It entails analysing historical financial reports to identify patterns and predict future performance. This evaluation is conducted through an examination of the organization's balance sheet, cash flows, and income statement. Financial statement analysis is employed by both internal and external stakeholders of a company to effectively run the company, make prudent investments, and compare it with other companies.

It is used by external stakeholders to have an idea of the business value and the general well-being and performance of an organisation, while internal users utilize it as a management tool. In the past few decades, financial constituents carry out this analysis manually, following the accounting procedures they are acquainted with. However, the evolution of technology has revolutionized the financial sector. Financial statement analysis can now be done with the use of Machine Learning (ML) however; Nigeria financial sector has not fully integrated this technology into some areas in the industry.

Machine learning is a branch of Artificial Intelligence (AI) that allows computers to acquire knowledge and improve their performance by analysing and interpreting data and to discover patterns that are not evident to humans. It has been used for many applications, such as stock market prediction, fraud detection, and financial statement analysis. The process of machine learning for financial statement analysis is done by developing models that learn from available data and apply the acquired knowledge to make decisions and predictions. ML algorithms have become increasingly popular in financial statement analysis and forecasting due to their ability to automate the process and provide more accurate predictions. ML algorithms can identify patterns in data that are not easily detected by humans. This has reduced the amount of time and effort required for financial statement analysis and forecasting. This AI has been of great benefit to so many sectors most especially the Nigerian Stock Exchange Market which is the focus of this research (Obadiaru et al., 2020).

With the largest economy in Africa, Nigeria is home to the primary stock exchange, the Nigerian Stock Exchange (NSE). This exchange has experienced significant growth in the past two decades, and its market capitalization has grown from US$15 billion in 2007 to US$3.4 trillion in 2022 (Lawal et al., 2021). With the increasing complexity of financial markets, investors and financial analysts are increasingly relying on machine learning techniques to analyse and forecast stock prices of companies that are listed on the NSE (Wei et al., 2017). Furthermore, machine learning algorithms can be employed to create predictive models for forecasting stock values.

Research Model

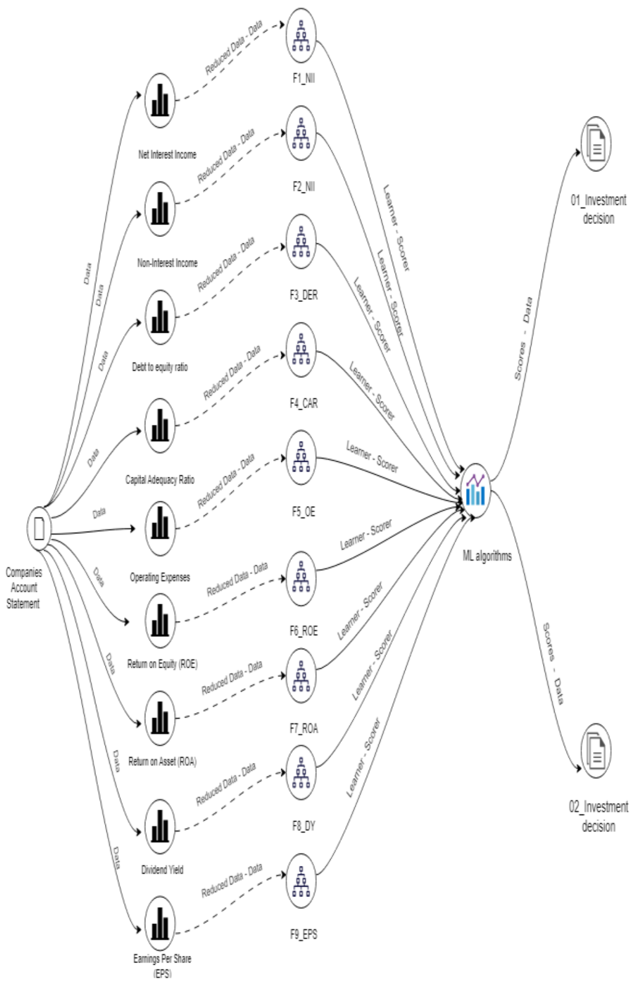

Figure 2.1 Research model

2. Literature Review

In order to establish a background for the theoretical framework of this study, this section contains the evaluation of previous studies on variables to consider in Financial Statement Analysis. Since GTCO Plc is a leading company listed on the Nigeria Stock Exchange Market and investors can only get their portfolios through the NSE, it is judicious to analyse the variables for Financial Statement Analysis in the Nigeria Stock Exchange Market.

Financial statement analysis has long been a cornerstone for evaluating firm performance, guiding investment decisions, and forecasting future trends. Traditionally, this analysis relied on accounting ratios and trend comparisons, but the evolution of financial markets has introduced greater complexity, requiring more advanced approaches such as machine learning (ML). This review synthesizes existing literature on financial performance variables, analytical methods, and the role of ML in enhancing financial forecasting, with a particular focus on the Nigerian Stock Exchange (NSE) and GTCO Plc.

Independent Variables in Financial Analysis

Macroeconomic indicators are central to financial performance assessment. Studies have shown that variables such as GDP, inflation, and interest rates significantly influence stock returns (Fama & French, 1988; Amadi, 2021). In Nigeria, exchange rate fluctuations (Lawal et al., 2017) and interest rate changes (Odo et al., 2019) have been directly linked to firm profitability and stock price volatility. Industry trends, regulatory reforms, and company size also matter: larger firms tend to demonstrate stronger resilience, while smaller firms are more sensitive to market shocks (Fasan & Gbadebo, 2017). Market capitalization is another determinant, with several Nigerian studies linking higher capitalization to better long-term performance (Olaoye & Eriki, 2019).

These variables act as external determinants that shape how firms perform and how their financial statements are interpreted. However, prior literature reveals a gap in integrating these indicators with predictive models that can capture nonlinear relationships — a challenge where ML techniques are particularly well-suited.

Dependent Variables in Financial Analysis

On the performance side, several dependent variables have been consistently highlighted:

Revenue and Profit Margin: Core indicators of firm health, with declining margins often signaling inefficiency or rising costs (Bustani et al., 2021).

Earnings Per Share (EPS): Widely used as a benchmark for profitability and valuation, with empirical evidence showing strong links between EPS and firm value in Nigeria (Adekunle et al., 2020; Inanga et al., 2016).

Debt-to-Equity Ratio: A key measure of leverage and financial risk. High leverage may reduce profitability, as seen in MTN Nigeria's financial restructuring (Adesina et al., 2017).

Return on Equity (ROE) and Return on Assets (ROA): Standard profitability ratios. Although extensively studied, sector-specific variations remain underexplored, particularly within Nigerian banking.

Dividend Yield and Stock Price: Critical for investor decision-making. Empirical studies show dividend policies can significantly influence stock valuations, though findings vary by sector (Majeed et al., 2017).

These indicators form the foundation of both traditional financial analysis and ML-based predictive models.

Methods of Financial Statement Analysis

Classical approaches include horizontal and vertical analysis, as well as ratio analysis. Horizontal analysis detects year-to-year trends in revenues, costs, and profits (Khalid et al., 2016), and has also been applied in fraud detection (Bhattacharya et al., 2002). Vertical analysis evaluates financial structure relative to a common base, aiding comparisons across firms and sectors (Raza et al., 2015). Ratio analysis provides diagnostic insights into liquidity, solvency, and efficiency, though studies caution that reliance on ratios alone may obscure broader trends (Karimi & Rahmani, 2017).

While effective, these approaches are limited in handling high-dimensional and nonlinear financial data — conditions prevalent in volatile markets such as Nigeria. This has spurred the integration of machine learning.

Sustainable Competitive Advantage Indicators

Beyond financial metrics, literature also highlights the role of sustainable competitive advantage (SCA). Factors such as innovation, brand strength, customer loyalty, and efficient cost structures underpin long-term performance (Purwanto & Purwanto, 2020). In the Nigerian context, financial institutions with diversified revenue streams and strong capital adequacy, such as GTCO Plc, have demonstrated resilience. However, empirical frameworks for quantitatively linking SCA to financial performance remain limited, offering scope for ML-driven models.

Machine Learning in Financial Analysis

The integration of ML has transformed financial statement analysis by offering predictive power and adaptability to complex datasets. A growing body of work has applied ML to stock price prediction, financial distress forecasting, and ratio analysis.

Algorithmic Applications: Logistic Regression, Random Forest, Support Vector Machines (SVM), Neural Networks, and ensemble methods such as stacking have been widely tested (Oyewola et al., 2019). Studies emphasize the superior accuracy of ensemble methods when compared to single-model approaches.

Validation Techniques: Cross-validation methods (K-fold, repeated K-fold) have been shown to improve robustness, particularly in emerging markets like Nigeria (Ogundunmade et al., 2022).

Domain Applications: In the Nigerian Stock Exchange, ML models have been applied to forecast stock returns (Akande & Olubiyi, 2019), assess credit risk (Altman & Sabato, 2005), and predict financial performance (Al-Nabulsi et al., 2020). These studies consistently show that ML outperforms traditional statistical approaches in predictive accuracy.

Despite these advances, gaps remain in tailoring ML models to firm-specific contexts, such as banking, where variables like interest income, capital adequacy, and operating efficiency carry unique significance. This thesis addresses such gaps by focusing on GTCO Plc, a leading Nigerian bank, to test the utility of ML algorithms against traditional financial analysis.

Synthesis and Relevance to Current Study

The reviewed literature reveals three main insights. First, macroeconomic and firm-specific variables exert significant influence on financial outcomes, yet their interactions are complex and often nonlinear. Second, while traditional methods provide useful descriptive insights, they fall short in predictive accuracy. Third, ML offers promising solutions by capturing hidden patterns and relationships, but its application in Nigerian banking remains underexplored.

This study builds on these gaps by applying ML algorithms specifically Logistic Regression, Random Forest, K-Nearest Neighbours, and Naïve Bayes to GTCO Plc's financial statements over 12 years. By comparing their predictive accuracy with traditional horizontal analysis, the research contributes to both theory and practice. It demonstrates how AI techniques can enhance financial forecasting, support investment decision-making, and ultimately identify sustainable competitive advantages within the Nigerian banking sector

3. Methodology

The mixed methodology was employed to achieve a full understanding of this investigation. This approach comprises both quantitative and qualitative research methods. This method facilitates a comprehensive analysis of the issue by taking into account both subjective and objective perspectives.

The quantitative method involves collection and analysing of numeric data from one of the companies listed on the Nigerian Stock Exchange market (NSE), GTCO Plc. An in-depth analysis of the company's financial statement of a period of 12 years was carried out using the traditional financial statement analysis method. The qualitative method was used in this thesis majorly for the literature review.

Previous studies on financial statement analysis and Machine Learning were thoroughly examined, finding gaps in the literature and posing questions for further studies. This method was also used to critically examine previous research on the independent and dependent variables of this subject. It also assessed what researchers have studied in respect to identifying the sustainable competitive advantage of companies listed on the NSE thereby addressing one of the objectives of this study. Finally, this method allowed me to evaluate the machine learning techniques and models performed in previous studies.

3.1. Sampling Method

A non-probability sampling method will be used to sample the thesis topic. The study utilizes purposive probability sampling technique to collect data from the financial statements of Guarantee Trust Bank Limited (GTCO Plc.), which is one of the firms listed on the Nigerian stock Exchange Market (NSE). This sampling technique was chosen because the data does not involve a random selection

3.2. Source of Data

Data will be gathered from GTCO Plc, one of the prominent banks in the Nigeria stock exchange market, using purposive sampling, a non-probability sampling method. In this instance, GTCO was chosen as a data source because it has the largest financial statement of banks listed on the stock exchange market. GTCO's annual reports encompasses of the metrics and ratios needed to analyze the data. The data accessible from the company enable us determine the sustainability of the firm.

3.3. Data Collection

To collect data for this study, a secondary data collection method was used. The data was collected from Guarantee Trust Bank Limited (GTCO Plc.**)** financial statement over the period of 12 years. Using a 12-year sample size of data from the list of GTCO annual report, this data would be used to develop a model and create a trend that the machine learning can follow in order to forecast the performance of the company.

3.4. Data Analysis

The Annual Reports of Guaranty Trust Bank Limited from 2009 to 2022 was analyzed technically using the horizontal financial statement analysis method and some ratios like Return to Equity (ROE), Return on Asset (ROA), Net Interest Income, Non-Interest Income, Debt to equity ratio, Capital Adequacy Ratio, Operating Expenses, Dividend Yield, etc.

3.4.1. Horizontal financial statement analysis

Horizontal analysis is used to analyze historical data, such as ratios and line items, from several accounting periods. Horizontal analysis involves making comparisons using either absolute or percentage comparisons. Absolute comparisons include comparing data directly, whereas comparing percentages describe the numbers as a percentage of the total for the previous and succeeding periods, with 100% being the baseline value. This is sometimes referred to as the base-year analysis.

Return on Equity (ROE): Return on Equity (ROE) is a financial metric used to Return on Equity (ROE) is a financial metric used to assess a company's profitability. It is calculated by dividing the company's net profit by its shareholder equity, indicating how much profit is generated for each dollar of equity. ROE is a measure of a firm's ability to create returns from the investments made by its investors. Based on the analysis conducted in this study, a higher return on equity (ROE) signifies that the bank is earning greater profits from the capital invested by shareholders.

$$ROE = \frac{Net\ Income}{shareholder\ Equity}$$

Return on Assets (ROA): Return on assets (ROA) is a financial metric that quantifies the profitability of a firm in relation to its overall assets. The metric quantifies the degree to which a corporation is effectively utilizing its assets to produce profits. A substantial proportion of this ratio signifies the company's management's effectiveness in utilizing its balance sheet to produce profits. Return on assets (ROA) is computed as:

$$ROA = \frac{Net\ Income}{Total\ Assets}$$

Net Interest Income: Net interest income refers to the net amount of money a financial organization gets from the interest it receives on its assets, such as loans and securities, less the interest it pays on its obligations, such as deposits and borrowings. It is a key component of a financial institution's earnings. A bank or institution with a high NIL indicates that it has a strong core business

$$ROE = Interest\ Income - Interest\ Expenses$$

Non-Interest Income: Non-interest income refers to the revenue that a financial institution earns from activities other than lending and borrowing money. It consists of commissions and fees received for services including insurance, brokerage, and investment management. If a bank has a diversified revenue stream, it helps reduce the bank's reliance on interest income.

$$ROE = Total\ Revenue - Net\ Interst\ Income$$

Debt to Equity Ratio: The debt-to-equity ratio expresses how much of a company's entire debt is divided by its total equity. The calculation involves dividing the sum of all liabilities by the amount of shareholder equity. The ratio is a gauge of how much of the company's funding is provided by debt as opposed to equity. The optimal debt-to-equity ratio (D/E Ratio) differs among industries, however it is generally recommended to not exceed 2.0. Industries that require a significant amount of capital, such as mining and manufacturing, typically exhibit a higher ratio. However, for our data, the ideal D/E ratio for a bank should be from 1.5 – 2.0

$$Debt\ to\ equity\ ratio = \frac{Total\ Liabilities}{Shareholder\ Equity}$$

Capital Adequacy Ratio: The capital adequacy ratio is a metric that assesses a bank's capacity to withstand losses that may result from its operational undertakings. The calculation involves dividing a bank's capital, which includes equity and retained earnings, by its risk-weighted assets. A greater capital adequacy ratio signifies that the bank possesses a robust financial position and is capable of withstanding bad economic situations. The Central Bank of Nigeria has established that all banks and banking groups with international authorization, as well as those categorized as Domestic Systemically Important Banks by the CBN, must have a minimum Pillar 1 regulatory Capital Adequacy Ratio of 15%. All other banks would be required to maintain a minimum Capital Adequacy Ratio (CAR) of 10%.

$$Caapital\ Adequacy\ Ratio = \frac{Capital}{Risk\ Weighted\ Assets}$$

Operating Expenses: Operating expenses refer to the expenditures incurred by a corporation for its routine business activities, including employee wages, rental fees, utility bills, and advertising costs. They exclude expenses associated with the manufacturing of products or services, such as the expenses for acquiring raw materials or labor.

$$Operating\ Expenses = \frac{Operating\ Expenses}{Total\ Revenue}$$

Dividend Yield: Dividend yield is a financial metric that quantifies the proportion of cash dividends distributed to shareholders in relation to the market value of a company's stock. The chairman of an organization will usually declare the year's dividend yield in the annual general meeting and it would be stated in the annual report. However, it can be ascertained by dividing the yearly dividend per share by the stock price per share.

$$\text{Dividend yield} = \frac{Dividend\ per\ share}{Share\ Price}$$

Earnings per Share (EPS): Earnings per share (EPS) is a financial metric used to gauge a company's profitability by calculating the net income made per share of its outstanding common stock. A higher EPS is a positive signal to investors, although decisions from this ratio is relative to each investor depending what the returns they expect on their investment. The earnings per share is usually stated in the firm's annual report, however, it can be calculated as follows;

$$\text{EPS} = \frac{Net\ Inocme - Dividend\ Paymnet}{Weighted\ Average\ of\ common\ share\ outstanding}$$

Stock Price: The stock price is the present market valuation of an individual share of a company's stock. The stock market's dynamics are determined by the interplay of supply and demand and reflects investors' perception of a company's future earnings potential, risk, and growth prospects.

3.4.2. Machine learning model and algorithm

The four major machine learning model that would be used for this thesis are Random Forest, Logistics Regression, KNN, and Naïve Bayes.

Random Forest: Random Forest is a versatile machine learning technique employed for tasks such as classification, regression, and feature selection. It is a technique in ensemble learning where many decision trees are created during training and the final prediction is determined by either selecting the most common class among the trees or calculating the average prediction of each tree. The Random Forest algorithm operates by generating decision trees on randomly sampled subsets of the data and variables. At each split, the algorithm selects the best variable among a random subset of variables to create a split, which leads to a more robust and accurate model that is less prone to over fitting. Due to its capacity to process complex data, manage noisy data, and provide profound clarifications on the importance of various financial ratios and metrics in forecasting financial performance indicators, it is extensively employed in financial statement analysis.

Logistic Regression: Logistic Regression is a statistical technique employed to solve binary classification problems. It quantifies the likelihood of a binary result (such as default or non-default) based on one or more independent variables (such as financial ratios or metrics). Logistic Regression operates by calculating the coefficients of the independent variables that optimize the likelihood of the observed data. The coefficients quantify the impact of each independent variable on the outcome and can be utilized to forecast the likelihood of the outcome based on the values of the independent variables. Logistic Regression is widely used in financial statement analysis for its simplicity, interpretability, and ability to handle both continuous and categorical data. It is often used to predict financial distress, bankruptcy, and other binary outcomes in corporate finance.

Naive Bayes: Naive Bayes is a probabilistic technique in machine learning that is specifically designed for solving classification problems. It relies on Bayes' theorem, which quantifies the likelihood of a hypothesis given the available evidence. Naive Bayes operates by estimating the conditional probability of each class given the input variables, assuming that the variables are independent of each other. This assumption is often unrealistic, hence the name "naive", but it still works well in practice for many classification problems, especially in the presence of high-dimensional and sparse data. Naive Bayes is widely used in financial statement analysis for its simplicity, speed, and accuracy in predicting financial distress, bankruptcy, and other classification problems in corporate finance.

K-Nearest Neighbor: The K-Nearest Neighbours (kNN) algorithm is a straightforward and adaptable supervised machine learning technique employed for classification and regression tasks. The algorithm is based on the idea of proximity, whereby an input data point is categorized by the predominant class of its k closest neighbors in the feature space. The selection of k influences the quantity of neighbors that are taken into account throughout the classification process. The algorithm's simplicity facilitates comprehension and implementation, but, its performance can be influenced by the selection of the distance measure and the value assigned to k. Moreover, kNN is a non-parametric approach, implying that it does not rely on assumptions regarding the underlying data distribution. This characteristic makes it well-suited for datasets that exhibit diversity.

4. Data Analysis

4.1. Interpretation of Data Analysis

The data we collected for this study encompasses of GTCO's annual report from the year 2009 to 2022, excluding 2010 and 2015. The horizontal financial statement analysis method was used to traditionally analyze the bank's financial statement and the results derived from this analysis are represented in the table below;

When analyzing financial statement most especially for investment purposes, there are certain metrics analysts look out for to determine the best investment. Investors that are not well informed of financial analysis have been educated to only look out for metrics or ratios like Earnings per Share, Net Income, Return on Assets (ROA), Return on Equity (ROE), and Dividend Yield etc. As important as these rations are, they are not only the factors that determines the bank's general state of health, which will aid in their ability to make wiser investment choices.

In determining the best investment decision, each of these metrics and ratios are evaluated in accordance with the general guidelines of the industry. Results gotten for each data are explicitly explain below.

Table (1): Traditional analysis of GTCO Annual Report 2009-2022.

Year	Net Interest Income	Non-Interest Income	Debt to Equity Ratio	Capital Adequacy Ratio	Operating Expenses	Return on Equity (ROE)	Return on Asset (ROA)	Dividend Yield	Earnings per share (EPS)	Stock Price
2009	73,468,110	40,808,407	4.4	25.99%	82%	12%	2.34%	1.5%	128 Kobo	50 kobo
2011	95,522,806	48,378,608	1.0	23.03%	63%	22%	3.34%	8.97%	164 Kobo	₦12.25
2012	123,098,741	44,199,867	4.6	24%	32%	30%	5.26%	9.87%	₦2.90	₦15.70
2013	127,857,215	49,285,953	4.8	22.27%	44.47%	25.95%	4.49%	7.97%	₦2.91	₦21.31
2014	128,698,830	69,022,777	4.8	21.40%	53.08%	25.28%	4.39%	6.80%	₦3.17	₦22.03
2016	171,027,957	139,337,380	4.5	19.79%	40.54%	26.59%	4.85%	8.16%	₦4.31	₦24.61
2017	217,649,619	75,794,761	3.8	25.50%	45.10%	27.60%	5.70%	10.21%	₦5.48	₦26.44
2018	188,441,907	98,521,378	4.3	23.4%	46.26%	32.61%	6.15%	7.61%	₦5.67	₦36.12
2019	189,318,029	109,749,294	4.1	20.66%	42.42%	28.90%	5.65%	9.30%	₦5.95	₦30.08
2020	208,932,503	123,485,182	4.8	19.55%	41.23%	25.36%	4.38%	13.16%	₦6.05	₦22.79
2021	0	8,829,354	4.41	23.83%	6.2%	6.02%	5.76%	11%	28 Kobo	₦27.83
2022	0	2,092,332	1.89	24.08%	8.17%	64%	54%	13%	₦3.01	₦24.38

4.1.1 Net interest income

In 2009, the net interest income for GTCO was ₦73,468,110 out of a total interest income interest of ₦110,889,700, this indicate a high interest income which means The bank is generating more revenue from its interest-earning assets than it is paying out on its interest-bearing liabilities. of ₦37,421,590 which is indicative of a wise investment. The financial statement in 2011 reported a net income interest of ₦95,522,806 reflecting a 30% increase from the previous year. The net interest rate increased the following year 2012, with 28.8%, the annual report recorded ₦123,098,741 as total Net interest income.

However, the net interest income didn't reflect much increase in 2013 and 2014, it grew slightly from ₦127,857,215 to ₦128,698,830 reflecting an increase of 3.9% and 0.66% respectively. This is due to the fact that the bank had little revenue on interest from loans and securities and they had little expenses to pay on their interest-bearing liabilities like deposits and borrowings. The net interest income picked up in 2016 by 32.89% with the total net income interest being ₦171,027,957. In 2017, the net interest income grew from ₦171,027,957 to ₦217,649,619 representing a 27.3% increase and the highest metric in our data set. The results gotten from 2018 to 2022 are represented below in the bar chart. There was a 0.46% increase within 2018 to 2019 and a 10.36% from 2019 to 2020. However, the financial statement of the bank in 2021 and 2022 did not reflect any net income interest.

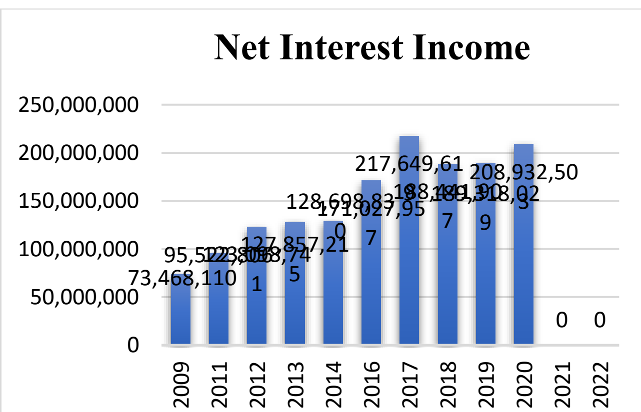

Figure 4.1: GTCO net income interest 2009 -2022

4.1.2 Non-interest income

As established earlier, Non-interest income are the revenue a financial institution earns from other activities than borrowing and lending money. These other source or revenue could be generated from fees like commission income, trade gains, insurance income, etc. The Non-Income Interest derived for this study are represented in the figure below.

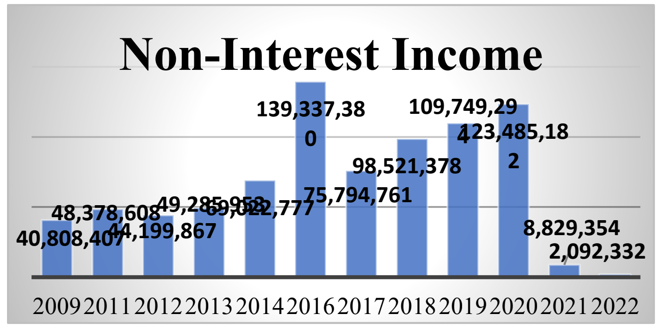

Figure 4.2: GTCO non-interest income 2009-2022

The bank's non-interest income amounted to ₦40,808,407 in 2009, and in 2011, their annual statement reported ₦48,378,608 as Non-interest income reflecting 18.6% increase. However, this metric reduced by 3.7% in 2012 and picked up again by 11.5% and 40% in 2013 and 2014 respectively. In 2016, the non-interest income increased from ₦69,022,777 in the previous year (2014) to ₦139,337,380 reflecting more than a 100% increase. This increase might not reveal the accurate picture of the data due to the year gap of 2015 in our data set, this year gap is one of the limitations faced in this study which will be discussed later on in this study.

This metric took a dive in 2017 with a decrease of 45.6% which means that the bank could not generate more alternative income like the previous year. In 2018, the bank generated ₦95,521378 revenues and ₦109,749,294 in 2019. The non-interest income grew by 12.5% generating an income of ₦123,485,182 in 2020. GTCO annual report in 2021 and 2022 reported a very low non income statement income of ₦8,829,354 and ₦2,092,332.

4.1.3 Debt to equity ratio

A company's debt-to-equity ratio quantifies the extent to which it is using its own resources to finance its operations instead of borrowing from others. Analysing a debt to equity ratio for investment purposes helps us identify companies that are highly leveraged and higher risk. The ideal D/E ratio for a bank should be from 1.5 – 2.0, any ratio higher that this means that the company is a highly leveraged.

The graph below represents the D/E Ratio of GTCO from 2009 - 2022. For this analysis the company started off with a high risk ratio of 4.4. Considering the banking industry ideal rate of 1.50 - 2.0, the 2009 ratio was a bad indication on the banks annual report that year. Moreover, since financial institutions borrow money to lend money, they tend to have higher D/E Ratio, but according to the Central Bank of Nigeria (CBN) the ideal D/E Ratio for financial sector should be 4%. This ratio measures the firm's financial leverage, i.e. for every amount of equity it shareholders own, the bank owes ₦4.4.

In 2011, the firm's annual report reflected a low D/E Ratio of 1% i.e., a debt of ₦1 for every amount of equity the bank shareholders own. This ratio is within the scope of the industry's acceptable rate, therefore the annual report for the year was not highly leveraged. From 2012 – 2021 the company's D/E Ratio was highly leveraged, the ratio fluctuated within 4.1 to 4.8 but with an exception in 2017. In 2017, the rate was 3.8 which is above the ideal ratio for financial institutions but below the set rate from the Central Bank of Nigeria (CBN). In this case, I decided to follow the general acceptable industry ratio to make a decision of D/E Ratio being high risk. However, 2022 D/E Ratio was 1.89 which indicated a low risk and good for making investment decision.

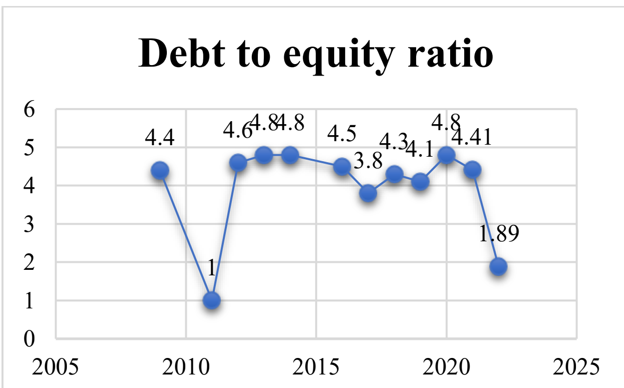

Figure 3.3: GTCO debt to equity ratio 2009-2022

4.1.4 Capital adequacy ratio

A capital adequacy ratio of 25.99% in 2019 denotes an excellent financial position of the bank. This denotes that GTCO has the ability to absorb losses with their capital if any arises. According to the Central bank of Nigeria (CBN), all Nigerian banks with an international authorisation in which GTCO has, is expected to have at least 15% capital adequacy ratio. The bank had more than the required rate in all the 12 years' data set for this study, which could be attractive to the investors because they are sure their investment would be safe to some extent if any crisis or liquidation arises. The figure below gives a graphical representation of the capital adequacy ratio for 12 years.

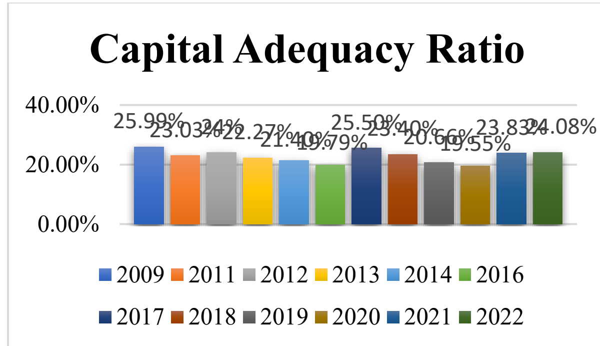

Figure 4.4: GTCO capital adequacy ratio 2009-2022

4.1.5 Operating expenses

Operating expenses are the financial metric that evaluates the ongoing operational activities of an organisation. A high operating expenses means that the company may have difficulty in generating profits. This can influence the decision of the investors, they will consider the negative effect this will have on the firm's profitability and also its cash flow. A high operating expense can have a negative impact on the firm's cash flow, it would be difficult for them to pay dividends to shareholders. GTCO's operating expenses from 2009-2022 is represented in the bar chart below:

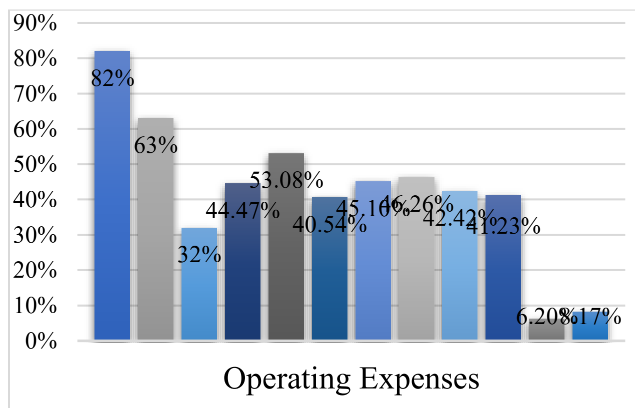

Figure 4.5: GTCO operating expenses 2009-2022

4.1.6 Return on Equity

Return on Equity is a financial metric that calculates the profit made per unit of shareholder stock in order to quantify the profitability of a business. ROE is a measure of a firm's capacity to create returns from the investments made by its stakeholders in an efficient manner. This metric is one of the most important indicators investors look out form while making investment decisions. A high return on equity (ROE) signifies that the company is efficiently producing greater profits from the capital invested by its investors. Although, ROE varies from sector to sector but a good return on equity rate should range from 15% - 20% and above.

In 2009, the Return on Equity for this year was 12% which does not indicate a good investment, it could mean that the company did not use the shareholders' investment efficiently.

ROE from 2011 to 2022 were above the 15% which will be very attractive for potential stakeholders.

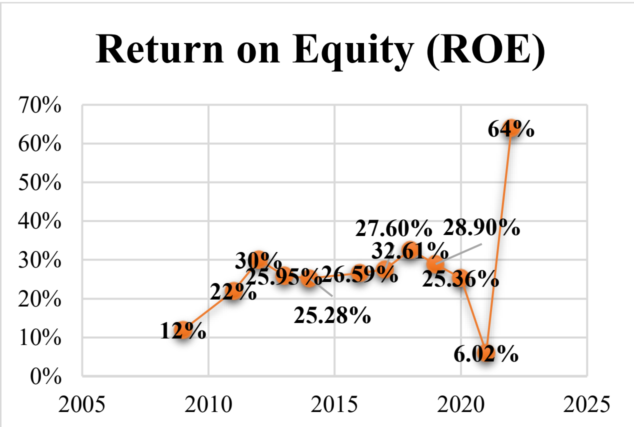

Figure 4.6: GTCO Return on Equity 2009-2022

4.1.7 Return on Asset

Return on Asset is an assessment of how well a business uses its earnings to make more profits. A high percentage of this ratio indicates how efficient the company's management handling its balance sheet to make profit. According to Investopedia, a ROA of more than 5% or more is considered good, while more than 20% is considered excellent. The graph below presents the Return on Equity of GTCO over a 12-year financial period.

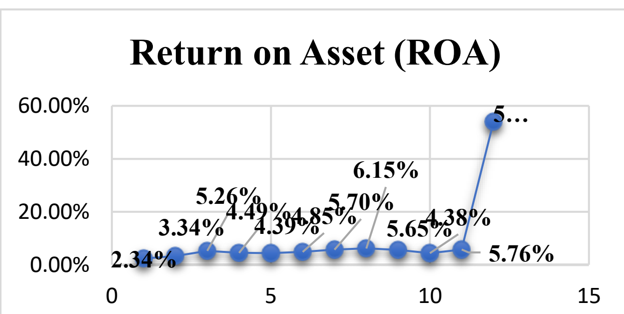

Figure 4.7: GTCO Return on Asset 2009-2022

4.1.8 Dividend yield

Dividend yield is a financial ratio that quantifies how much an investor would earn from an investment based only upon the dividend payments from the organization. This metric is one the most important things investor looks out for when making an investment decision. Investors who does not have adequate financial knowledge are taught to look out for this metric to make their decision when buying a stock. The chairman of an organization will usually declare the year's dividend yield in the annual general meeting and it would be stated in the annual report.

Some investors evaluate stocks based on only the dividend yield which is not advisable because a dividend can reduce or even be eliminate if the stock price declines. This proves the importance of this study. Having a trained machine learning program to accurate predict the best investment while carefully evaluating each metrics will eliminate bad investments decisions. The linear graph below shows the dividend yield of GTCO from 2009 – 2022.

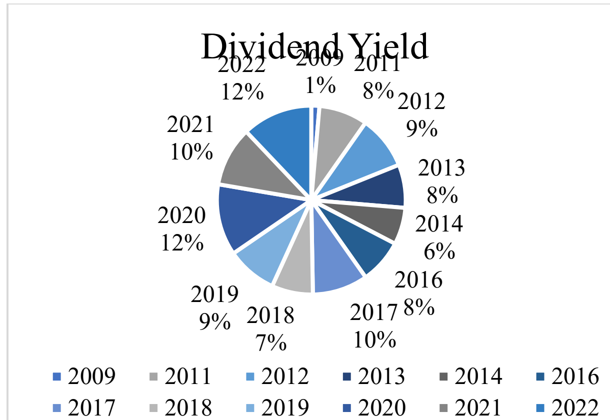

Figure 4.8: GTCO dividend yield 2009-2022

4.1.9 Earnings per Share

The Earnings per Share is a vital indicator for stockholders as it assesses the bank's profitability per share. A higher Earnings Per Share (EPS) signifies that the bank is making more money for every outstanding share of stock. For example, in 2009 on the graph below, the EPS was 128 Kobo, i.e. upon every outstanding share, 128 Kobo profit is earned. Although this matric is not enough to determine a firm's profitability, Investors are able to make well-informed decisions with the metric when compared with the EPS of other companies in the same industry.

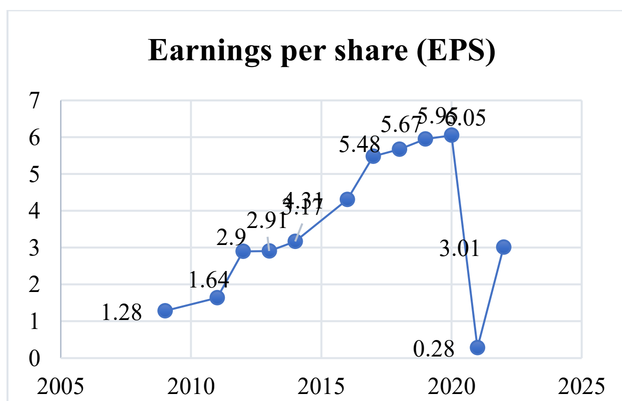

Figure 4.9: GTCO Earnings Per Share 2009-2022

4.1.10 Stock price

The stock price on the Nigerian Stock Exchange (NSE) represents the current market value of the bank's shares. A higher stock price shows that the market has confidence in the bank's future prospects and financial performance. GTCO's stock price experienced a steady increase from 2009-2018 and fluctuated from 2019-2022 which could be due to investors sentiment and market conditions.

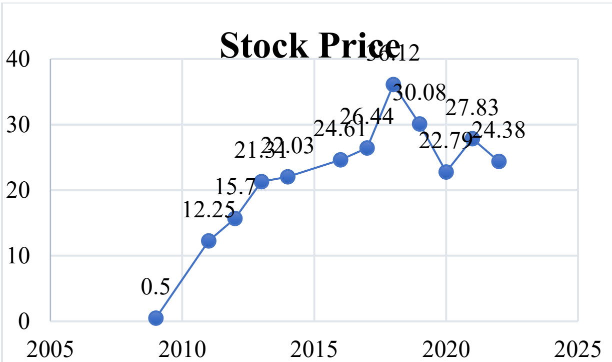

Figure 4.10: GTCO stock price in Naira

5. Conclusion and Discussion

5.1 Experiment and Result

This chapter presents the result of the experiment performed for this study. The primary objective of this study is to create a machine-learning model that will transform financial statement analysis from its traditional methods and to predict a company's financial health or profitability to enhance well informed investment decision. To create this model, this study used the supervised machine learning, which means that our model will have a baseline understanding of what the correct output should be.

In order to get a supervised dataset for our machine learning, it was necessary to first analyse the financial statement traditionally to train the machine to predict on its own. To achieve this, the table below represents my response to each of the metrics analysed in accordance to the financial industry overall acceptable rate. The last column title "Decision" is to give the machine learning algorithms the final decision on a year's financial report to either invest or not to. The tick sign (✓) indicates a positive sign and to invest while the sign (✕) means negative and not to invest. Training the machine with this data has to be done in a language the machine understands i.e. Binary language, so this data was inputted to the machine learning algorithms used in 1 and 0. 1 represents the positive (✓) sign while 0 for the negative (✕) sign.

Table (2): Decisions for traditional financial statement analysis

Year	Net Interest Income	Non-Interest Income	Debt to Equity Ratio	Capital Adequacy Ratio	Operating Expenses	Return on Equity (ROE)	Return on Asset (ROA)	Dividend Yield	Earnings per share (EPS)	Decision
2009	✓	✓	✕	✓	✕	✕	✕	✕	✓	✕
2011	✓	✓	✓	✓	✓	✓	✕	✓	✓	✓
2012	✓	✓	✕	✓	✓	✓	✓	✓	✕	✓
2013	✓	✓	✕	✓	✓	✓	✕	✓	✕	✓
2014	✓	✓	✕	✓	✓	✓	✕	✕	✓	✓
2016	✓	✓	✕	✓	✓	✓	✕	✓	✓	✓
2017	✓	✓	✕	✓	✓	✓	✓	✓	✓	✓
2018	✓	✓	✕	✓	✓	✓	✓	✓	✓	✓
2019	✓	✓	✕	✓	✓	✓	✓	✓	✓	✓
2020	✓	✓	✕	✓	✓	✓	✕	✓	✓	✓
2021	✕	✕	✕	✓	✓	✕	✓	✓	✕	✕
2022	✕	✕	✓	✓	✓	✓	✕	✓	✕	✕

The image below is the representation of the prediction of this study result from the four ML algorithms used. Each of this algorithms were running on every instances on our data i.e. trained to predict the result of all the financial metrics used for this study. On the extreme left of this image is the ML algorithms and the accuracy of their readings. As mentioned earlier, binary numbers were used to the results ascertained from the traditional financial analysis performed. The binary number "1" denote a positive result of "To Invest" and the binary number "0" symbolizes a negative result of "Not to Invest". The first four columns represent the accuracy of reading of Random Forest, kNN, Logistic Regression, and Naïve Bayes. In 2009, the final decision made on that year's analysis was not to invest i.e. 0 in binary. Random forest predicts to invest i.e. "1" with a degree of accuracy of 0.57, K nearest neighbor predicted to invest with 0.80 degree of accuracy. However, Logistic Regression predicted not to invest with the degree of accuracy of 0.58, while Naïve Bayes also predicted not to invest with 0.86 degree of accuracy. This algorithm both got the prediction correctly but the Naïve Bayes has the highest rate of prediction accuracy. For 2011, the decision on investment was to invest, all the algorithms predicted correctly for this year. This prediction was carried out through all instances in the dataset of GTCO, the algorithms predicted the result accurately with few confusion matrixes. The result gotten from this experiment proofs that the machine learning model was learning well and it supports the hypothesis that there is a positive relationship between the adoption of machine learning technique level of accuracy of financial statement analysis and forecast in the Nigerian stock exchange market.

Additionally, there are five machine learning metric to examine the level prediction accuracy of our algorithms.

1. **AUC (Area under the ROC Curve):** AUC is a common indicator used to estimate the discriminatory ability of a classification model. It measures the area under the Receiver Operating Characteristic (ROC) curve. A perfect AUC score is 1.000, which means the model has a perfect ability to differentiate between the classes. An AUC of 1.000 indicates that the model's predictions are completely accurate. As show in the image below, the AUC outputs for all model in this study were 1, which means all our ML model were able to distinguish between the classes and predict accurately.
2. **CA (Classification Accuracy):** Classification accuracy quantifies the percentage of occurrences properly classified out of all instances. In this case, kNN has a classification accuracy of 0.750 which means it correctly classified approximately 75% of the instances, while Random Forest correctly classified all instances by approximately 91.7%.
3. **F1 Score:** The F1 score is the harmonic mean of precision and recall. It strikes a balance between recall (the number of actual positive cases that were accurately predicted) and accuracy (the number of predicted positive instances that were truly positive). With an F1 score of 0.911, Random Forest exhibits a strong balance between recall and precision.
4. **Precision**: The ratio of accurately predicted positive instances to all expected positive instances is known as precision. It assesses how well the model avoids generating false positive results. A precision of 0.925 means that 92.5% of instances predicted as positive were actually positive.
5. **Recall:** Recall is a metric that quantifies the proportion of accurately predicted positive cases to all actual positive instances. It is sometimes referred to as sensitivity or true positive rate. With a recall of 0.917, the model was able to accurately identify 91.7 percent of the real positive events.

The figure below represents the rate of accuracy for each metrics, each bar on the chart shows the how accurately the ML algorithms predicted the result. The major limitation in the study is limited availability of data, due to this we could not get a wider range of result for his experiment, however it does not depute purpose of the study. The chart below supports the hypothesis that the accuracy of financial statement analysis and forecasting using machine learning algorithms is influenced by the type of machine learning algorithm used.

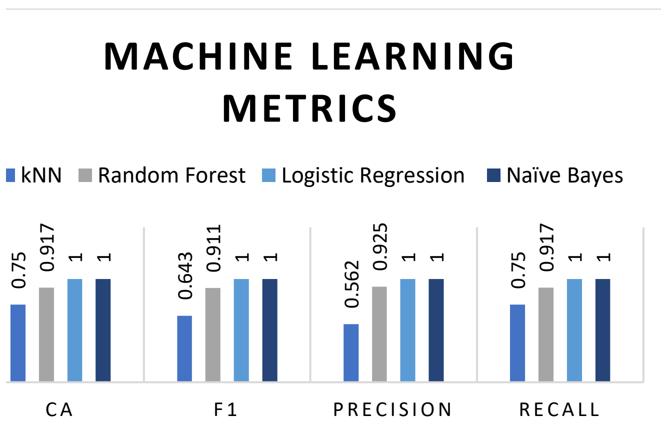

Figure 4.11: ML metrics rate of accuracy

Model: 4 models

- Random Forest
- kNN
- Logistic Regression
- Naive Bayes

Showing probabilities for all classes that appear in the data

Data & Predictions

	Random Forest	kNN	Logistic Regression	Naive Bayes	Decision	Net Interest Income	Non-Interest Income	Debt to equity ratio	Capital Adequacy Ratio	Operating Expenses	Return on Equity (ROE)	Return on Asset (ROA)	Dividend Yield	Earnings per share (EPS)
1	0.43 : 0.57 → 1	0.20 : 0.80 → 1	0.58 : 0.42 → 0	0.86 : 0.14 → 0	0	1	1	0	1	0	0	0	0	1
2	0.00 : 1.00 → 1	0.00 : 1.00 → 1	0.13 : 0.87 → 1	0.06 : 0.94 → 1	1	1	1	1	1	1	1	0	1	1
3	0.00 : 1.00 → 1	0.00 : 1.00 → 1	0.12 : 0.88 → 1	0.06 : 0.94 → 1	1	1	1	0	1	1	1	1	1	0
4	0.00 : 1.00 → 1	0.00 : 1.00 → 1	0.15 : 0.85 → 1	0.08 : 0.92 → 1	1	1	1	0	1	1	1	0	1	0
5	0.03 : 0.97 → 1	0.00 : 1.00 → 1	0.15 : 0.85 → 1	0.06 : 0.94 → 1	1	1	1	0	1	1	1	0	0	1
6	0.00 : 1.00 → 1	0.00 : 1.00 → 1	0.09 : 0.91 → 1	0.02 : 0.98 → 1	1	1	1	0	1	1	1	0	1	1
7	0.00 : 1.00 → 1	0.00 : 1.00 → 1	0.07 : 0.93 → 1	0.02 : 0.98 → 1	1	1	1	0	1	1	1	1	1	1
8	0.00 : 1.00 → 1	0.00 : 1.00 → 1	0.07 : 0.93 → 1	0.02 : 0.98 → 1	1	1	1	0	1	1	1	1	1	1
9	0.00 : 1.00 → 1	0.00 : 1.00 → 1	0.07 : 0.93 → 1	0.02 : 0.98 → 1	1	1	1	0	1	1	1	1	1	1
10	0.00 : 1.00 → 1	0.00 : 1.00 → 1	0.09 : 0.91 → 1	0.02 : 0.98 → 1	1	1	1	0	1	1	1	0	1	1
11	0.90 : 0.10 → 0	0.40 : 0.60 → 1	0.79 : 0.21 → 0	1.00 : 0.00 → 0	0	0	0	0	1	1	0	1	1	0
12	0.90 : 0.10 → 0	0.40 : 0.60 → 1	0.68 : 0.32 → 0	0.98 : 0.02 → 0	0	0	0	1	1	1	1	0	1	0

Scores

Figure 4.12 ML algorithms prediction of result

Target class: (Average over classes)

Model	**AUC**	**CA**	**F1**	**Precision**	**Recall**
kNN	1.000	0.750	0.643	0.562	0.750
Random Forest	1.000	0.917	0.911	0.925	0.917
Logistic Regression	1.000	1.000	1.000	1.000	1.000
Naive Bayes	1.000	1.000	1.000	1.000	1.000

Confusion Matrix Tue Aug 08 23, 12:47:51

Confusion matrix for Naive Bayes (showing number of instances)

		Predicted		
		0	**1**	**∑**
Actual	**0**	3	0	**3**
	1	0	9	**9**
	∑	**3**	**9**	**12**

Confusion Matrix Tue Aug 08 23, 12:48:18

Confusion matrix for Random Forest (showing number of instances)

		Predicted		
		0	**1**	**∑**
Actual	**0**	2	1	**3**
	1	0	9	**9**
	∑	**2**	**10**	**12**

Confusion Matrix Tue Aug 08 23, 12:48:22

Confusion matrix for Random Forest (showing number of instances)

		Predicted		
		0	**1**	**∑**
Actual	**0**	2	1	**3**
	1	0	9	**9**
	∑	**2**	**10**	**12**

Confusion Matrix Tue Aug 08 23, 12:48:27

Confusion matrix for kNN (showing number of instances)

		Predicted		
		0	**1**	**∑**
Actual	**0**	0	3	**3**
	1	0	9	**9**
	∑		**12**	**12**

Confusion Matrix Tue Aug 08 23, 12:48:31

Confusion matrix for Logistic Regression (showing number of instances)

		Predicted		
		0	**1**	**∑**
Actual	**0**	3	0	**3**
	1	0	9	**9**
	∑	**3**	**9**	**12**

Figure 4.14 Confusion matrix of ML algorithms

Confusion Matrix

The result above represent the confusion matrix which examins the rate of false-positive (the number of samples that are false but the ML model categorised as positive) and false-negative (the number of samples that are false but the ML model categorised as negative) (Kordi Ghasrodashti, Helfroush & Danyali, 2017) . For example Naïve Bayes did not have a condusion matrix, the model predicted all 12 instances accirately – 3 “Not to invest” prediction and 9 “To invest” prediction. This result corresponds with financial anlysis carried our traditionally.

Random Forest had one false-positive confusion matrix, the model predicted an instance which was “Not invest” as “To invest”. kNN had three false-positive confusion matrix, it predicted 3 instances out of 12 inaccurately which indicates that it is not really a good classification model. However, Logistic Rgression also had a perfect prediction.

It classified and predicted all instances accurately. In this instance, Naïve Bayes and Logistic Regression were the models with the perfect classification, the two other models has very little false-positives and this indicated that all our models could classify accurately. This reject the second hypothesis of this study which state that there is a significant difference between the performance of traditional financial statement analysis methods and machine learning algorithms in forecasting stock prices in the Nigerian stock exchange market.

Model	AUC	CA	F1	Precision	Recall
kNN	1.000	0.750	0.643	0.562	0.750
Random Forest	1.000	0.917	0.911	0.925	0.917
Logistic Regression	1.000	1.000	1.000	1.000	1.000
Naive Bayes	1.000	1.000	1.000	1.000	1.000

Rank Tue Aug 08 23, 13:38:14

Input

Features: Net Interest Income, Non-Interest Income, Debt to equity ratio, Capital Adequacy Ratio, Operating Expenses, Return on Equity (ROE), Return on Asset (ROA), Dividend Yield, Earnings per share (EPS)
Target: Decision

Ranks

		#	Info. gain	Gain ratio	Gini	χ^2	ReliefF	FCBF
1	Net Interest Income	2.0	0.4204484631347318	0.6468214774323336	0.22500000000000006	1.2	0.312	0.00013554029979531828
2	Non-Interest Income	2.0	0.4204484631347318	0.6468214774323336	0.22500000000000006	1.2	0.312	1.3554029979531828
3	Return on Equity (ROE)	2.0	0.4204484631347318	0.6468214774323336	0.22500000000000006	1.2	0.272	1.3554029979531828
4	Operating Expenses	2.0	0.1842428917900113	0.44522810430465803	0.10227272727272735	0.2727272727272727	0.1	4.301679078078116e-05
5	Earnings per share (EPS)	2.0	0.11556849565940197	0.12585105079823808	0.0625	0.6666666666666666	0.096	1.5425206156753393e-05
6	Debt to equity ratio	2.0	0.04300471205299761	0.0661588133282302	0.025000000000000133	0.6666666666666666		6.253906735093569e-06
7	Dividend Yield	2.0	0.04300471205299761	0.0661588133282302	0.025000000000000133	0.13333333333333333	0.012	6.253906735093569e-06
8	Return on Asset (ROA)	2.0	0.006987753258863694	0.007131315506727024	0.0035714285714286143	0.06666666666666667	-0.108	0.007863904808241394
9	Capital Adequacy Ratio	1.0				nan		

Output

Features: Net Interest Income, Non-Interest Income, Return on Equity (ROE), Operating Expenses, Earnings per share (EPS)
Target: Decision

5.2 Conclusion

This study has delved into the realm of machine learning for financial statement analysis and forecast within the context of the Nigerian Stock Exchange market, focusing on GTCO Plc. The overarching objective of this thesis was threefold: firstly, to predict the financial health and profitability of the company, thereby facilitating well-informed and data-driven investment decisions; secondly, to lower investors' risk of financial loss on the Nigerian Stock Exchange; and lastly, to progress machine learning's use in finance by creating a model that can replace traditional methods for financial statement analysis.

To achieve these objectives, we began by employing traditional methods of financial statement analysis, relying on the comprehensive annual reports of GTCO Plc. This initial analysis provided valuable insights into the company's financial standing, which served as the foundation for the subsequent development and evaluation of machine learning models.

The four machine learning algorithms chosen for this thesis Random Forest, K-Nearest Neighbor, Logistics Regression, and Naïve Bayes were systematically applied and assessed. Each algorithm brought its own strengths and limitations to the table, contributing to a well-rounded understanding of their efficacy in predicting financial outcomes. The results were promising, with the models demonstrating a capacity to offer valuable predictions regarding GTCO Plc's financial health and profitability.

The predictive power of these models is crucial in enhancing investment decision-making processes. Investors can now leverage these machine learning models to make informed choices, thus mitigating the risks associated with financial investments in the dynamic Nigerian Stock Exchange market. The reduction of financial losses is particularly significant in a market as dynamic and complex as Nigeria's, where economic and geopolitical factors can impact stock prices swiftly and unpredictably.

Furthermore, the successful development and adoption of machine learning models in this study enhances the field of financial analysis as a whole. By showcasing the effectiveness of these models in predicting financial outcomes, we advocate for the integration of machine learning techniques into financial analysis practices. This advancement stands to revolutionize traditional methods and bring about a new era of efficiency, accuracy, and adaptability in financial decision-making.

In essence, this thesis serves as a testament to the potential of machine learning in financial statement analysis and forecasting within the Nigerian Stock Exchange market. The application of these models not only benefits investors by providing valuable insights and reducing risks but also contributes to the ongoing evolution of financial analysis methodologies. As technology continues to progress, embracing machine learning in finance becomes imperative for staying ahead in an increasingly competitive and dynamic market environment. The findings of this study, centered on GTCO Plc, provide a foundation for future research and implementation of machine learning in financial markets, clearing the path for a more informed and data-driven method of making investing decisions.

5.3 Research Limitation

Despite the comprehensive nature of this study, it is imperative to acknowledge and discuss the limitations encountered during the research process. One significant constraint was the limited availability of data, as the analysis was confined to the annual reports of Guaranty Trust Bank Plc spanning only 12 years. Ideally, for the development and validation of robust machine learning models, a more extensive dataset would have been preferred, ranging from 50 to 100 years of annual reports.

The primary limitation arising from the restricted dataset is the potential impact on the generalizability of the machine learning models. A larger dataset allows for a more diverse representation of economic cycles, market conditions, and unforeseen events, facilitating the creation of models that can better adapt to a broader range of scenarios. The limited temporal scope of the dataset may lead to a model that is more tailored to the specific circumstances present during the analyzed years, potentially compromising its predictive power in different economic environments.

Furthermore, a larger dataset would have enabled a more granular examination of the machine learning algorithms' performance over time, providing insights into their adaptability and effectiveness in varying market conditions. The scarcity of data, particularly in the context of financial markets known for their volatility and susceptibility to external influences, could result in models that may not adequately capture the complexity and nuances of the Nigerian Stock Exchange market.

Additionally, the limited dataset could impact the robustness of statistical analyses and hypothesis testing conducted in the study. A bigger number of samples often strengthens the statistical power of findings, enhancing the reliability and validity of the results. The reduced sample size may limit the generalizability of the study's conclusions and may not fully represent the broader population of companies in the Nigerian Stock Exchange.

Despite these limitations, it is essential to recognize that the availability of historical financial data, especially for publicly traded companies, is often constrained by factors such as disclosure practices, data accessibility, and the duration of a company's existence. Researchers in the financial domain frequently encounter challenges related to data availability and completeness, and acknowledging these limitations is crucial for a transparent and accurate interpretation of the study's outcomes.

In conclusion, while the constraints related to the limited dataset pose challenges to the study's generalizability and the robustness of machine learning models, this acknowledgment serves as a foundation for future research. Addressing these limitations requires efforts to access more extensive and varied datasets, ensuring that subsequent studies can build upon and refine the models developed in this research, thereby advancing the adoption of machine learning in financial statement analysis and forecasting in the Nigerian Stock Exchange market.

5.4 Recommendation for Future Research

This study has proven the possibility of incorporating machine learning in financial statement analysis in the Nigerian stock exchange market (NSE) and based on the results gotten, here are some recommendations for future studies.

- To overcome the limitation of a relatively small dataset employed in this study, future research should focus on accessing a more extensive and diverse set of financial data. Collaboration with regulatory bodies, financial institutions, and other stakeholders could facilitate the compilation of a comprehensive dataset covering a more extended period and a broader range of companies. This expansion would aid to the development of more robust machine learning models with increased generalizability.

- Considering the growing interest in democratizing finance, future research could focus on developing user-friendly interfaces that enable investors, even those without extensive technical expertise, to leverage machine learning insights for their investment decisions. This would contribute to the accessibility and adoption of advanced financial analysis tools

Research on the Nigerian Stock Exchange should be broadened to include international comparative studies, as this would provide important information about how well machine learning models apply to various financial markets and regulatory contexts.

By addressing these areas in future studies, scholars can further advance the understanding of machine learning applications in financial statement analysis and forecasting, contributing to the development of more accurate, robust, and ethically sound tools for informed investment decision-making.

6. References

Adebiyi, O. (2020). Corporate Governance and the Nigeria Stock Exchange Market. *African Journal of Corporate Governance*, 1(1), 1-14. https://doi.org/10.11648/j.ajcg.20200101.11

Adegbite, O. S., Adebisi, O. F., & Adegbite, E. A. (2019). Exchange rate volatility and financial performance of commercial banks in Nigeria: Evidence from Guaranty Trust Bank Plc. *European Journal of Accounting, Auditing and Finance Research,* 7(6), 32-42.

Adegbaju, A., & Adaramola, A. (2018). Financial ratios and corporate performance: Evidence from Nigerian banking sector. *Journal of Financial Reporting and Accounting*, 16(4), 661-678.

Adegbite, E., & Adebola, S. (2015). Corporate governance and corporate finance in Nigeria. Corporate Governance: *The International Journal of Business in Society*, 15(3), 369-386.

Adekunle, S. O., Oguntuase, A. M., & Akanbi, A. M. (2020). Empirical Analysis of Earnings per Share and Return on Assets of Firms Listed on Nigerian Stock Exchange. *Journal of Accounting and Financial Management*, 6(1), 10-17.

Adesina, O. T., Azeez, O. R., & Adesina, A. O. (2017). The effect of debt to equity ratio on firm's financial performance. *International Journal of Management, Accounting and Economics*, 4(10), 690-702.

Adesina, K., Abdulrahman, M. D., & Olayemi, J. K. (2018). Predicting sustainable competitive advantage of Nigerian firms: A machine learning approach. *Cogent Business & Management*, 5(1), 1444094. doi: 10.1080/23311975.2018.1444094

Adeyemi, O. D., Olugbara, O. O., & Oyelade, O. J. (2020). Industry trends and financial performance prediction of selected companies in the Nigerian Stock Exchange using machine learning. *Journal of Computational Science*, 44, 101175.

Afolabi, R. O., & Adefila, J. O. (2018). Financial statement analysis and investors' decisions in Nigeria. *International Journal of Business and Management,* 13(8), 99-107.

Ahmadi, M., Kouhizadeh, M., & Khosravi, B. (2020). The impact of firm size on the financial performance prediction using data mining techniques. *Journal of Accounting and Management Information Systems*, 19(3), 465-484.

Ainsworth, S., Sadovskaya, I., Vinogradov, E., Courtin, P., Guérardel, Y., Mahony, J., & Sinderen, D. v. (2014). Differences in lactococcal cell wall polysaccharide structure are major determining factors in bacteriophage sensitivity. *Bio*, 5(3). https://doi.org/10.1128/mbio.00880-14

Alao, O. R., & Okunlola, J. T. (2020). The impact of industry trends on the financial performance of selected manufacturing firms in Nigeria. *International*

Journal of Management, Innovation & Entrepreneurial Research, 6(1), 21-34.
Al-Kilidar, H. M., Albattah, H. M., Alkhalafat, F. M., Alabdulwahhab, F. M., & Al-Matrouk, M. A. (2019). Employee satisfaction prediction in banking sector using machine learning techniques. *International Journal of Advanced Computer Science and Applications,* 10(10), 312-320. doi: 10.14569/IJACSA.2019.0101043
Al-Nabulsi, A. A., Alfaqawi, M. A., Alsmadi, M. K., & Al-Sarayreh, H. A. (2020). Predicting financial performance of companies listed in Amman Stock Exchange using machine learning techniques. *Sustainability,* 12(7), 2711. doi: 10.3390/su12072711
Al-Malkawi, H. A., & Masadeh, R. (2021). Predicting the financial performance of companies listed on the Amman Stock Exchange: A machine learning approach. *Journal of Finance and Investment Analysis,* 10(2), 66-82.
Altman, E. I. (1968). Financial ratios, discriminant analysis, and the prediction of corporate bankruptcy. *The Journal of Finance,* 23(4), 589-609.
Altman, E. I., & Sabato, G. (2005). *Modelling credit risk for SMEs: Evidence from the US market. Abacus,* 41(3), 257-268.
Amadi, V. C. (2021). The Impact of GDP on Nigerian Stock Market Returns. *Journal of Economics and Sustainable Development,* 12 (5), 64-71
Akande, O. J., & Olubiyi, O. A. (2019). Predicting stock prices of Nigerian banks: A machine learning approach. *Journal of Financial Data Science*, 1(1), 106-116. doi: 10.3905/jfds.2019.1.1.106
Akinniyi, O. D., Akinola, O. A., Aluko, O. O., & Oloyede, O. O. (2021). Modeling stock market returns in Nigeria using machine learning algorithms. *Journal of Financial Reporting and Accounting,* 19(1), 43-62.
Akpan, U. F., & Udoma, J. A. (2016). The relationship between inflation rate and stock prices in Nigeria. *Journal of Applied Economics and Business Research,* 6(3), 93-106.
Ayinde, T. O., & Oladipo, O. A. (2020). Machine learning approach for financial statement analysis: An application in the Nigerian stock exchange. *Journal of Financial Reporting and Accounting.*
Baker, M., & Wurgler, J. (2007). Investor sentiment in the stock market. *Journal of Economic Perspectives,* 21(2), 129-151.
Beaver, W. H. (1968). Market prices, financial ratios, and the prediction of failure. *Journal of Accounting Research,* 6, 179-192.
Berger, A. N., & DeYoung, R. (1997). Problem loans and cost efficiency in commercial banks. *Journal of Banking & Finance*, 21(6), 849-870.
Bhattacharya, U., Graham, J. R., & Harvey, C. R. (2002). The effects of financial statement credibility on judgmental forecasts of future earnings. *The Accounting Review*, 77(Supplement), 115-142.
Bustani, B., Kurniaty, K., & Widyanti, R. (2021). The effect of earning per share, price to book value, dividend payout ratio, and net profit margin on the stock price in indonesia stock exchange. *Jurnal Maksipreneur: Manajemen, Koperasi, Dan Entrepreneurship,* 11(1), 1. https://doi.org/10.30588/jmp.v11i1.810
Chen, M., Cheng, S., & Hwang, Y. (2005). An empirical investigation of the relationship between intellectual capital and firms' market value and financial performance. *Journal of Intellectual Capital,* 6(2), 159-176. https://doi.org/10.1108/14691930510592771.
Chinwudu, C. F. and Nwanna, I. O (2016). The effect of financial deepening on economic growth in Nigeria (1985 -2014). *IOSR Journal of Economics and Finance*, 07(04), 11-28. https://doi.org/10.9790/5933-0704011128.
Clemes, M. D., Gan, C., & Zhang, D. (2010). Customer switching behaviour in the Chinese retail banking industry. *International Journal of Bank Marketing*, 28(7), 519-546. https://doi.org/10.1108/02652321011085185.
Economic Indicator; Definition and How to interpret. https://www.investopedia.com/terms/e/economic_indicator.asp
Elshandidy, T., Hassanein, A., & Gaber, M. (2017). The relevance of dividend yield in stock return: Evidence from the UK stock market. *Research in International Business and Finance*, 42, 1312-1327.
Fasan, O. E., & Gbadebo, O. J. (2017). The impact of company size on financial performance: Evidence from selected companies in the Nigerian Stock Exchange. *Journal of Accounting and Financial Management*, 3(1), 1-13.
Fama, E. F., & French, K. R. (1988). Dividend yields and expected stock returns. *Journal of Financial Economics*, 22(1), 3-25.
Fong, N. M., Fang, Z., & Luo, X. (2015). Geo-conquesting: competitive locational targeting of mobile promotions. *Journal of Marketing Research*, 52(5), 726-735. https://doi.org/10.1509/jmr.14.0229
Goudreau, J. (2021). Machine Learning Algorithms in Finance. *Investopedia.* https://www.investopedia.com/terms/m/machine-learning-algorithms-finance.asp
Hamid, I. (1907). Les musulmans francais du nord de l'afrique. *Bulletin of the American Geographical Society*, 39(12), 758. https://doi.org/10.2307/198579
Hou, Y., Ni, Z., & Xu, X. (2020). Industry trends, regulatory changes, and stock price synchronicity: Evidence from China. *Emerging Markets Finance and Trade*, 56(1), 1-12.
Inanga, E. L., Ogbulu, O. M., & Ogoun, E. N. (2016). The relationship between earnings per share and market

value of firms listed in the Nigerian Stock Exchange. *International Journal of Finance and Accounting*, 5(6), 354-358.

Investopedia. What is considered a good ROA https://www.investopedia.com/terms/r/returnonassets.asp#:~:text=What%20Is%20Considered%20a%20Good,firms%20in%20the%20same%20sector.

Iyoha, F. O., Adebiyi, A. A., Adebiyi, M. A., & Olugbara, O. O. (2019). Pattern analysis of trading activities in the Nigerian stock exchange. *Heliyon,* 5(5), e01766.

Jalaludin, N. H., & Ibrahim, M. H. (2014). Predicting bank performance using financial statement analysis: A case of Malaysian commercial banks. *Procedia - Social and Behavioral Sciences*, 145, 171-178.

Jiménez, G., Ongena, S., Peydró, J. L., & Saurina, J. (2014). Hazardous times for monetary policy: What do twenty-three million bank loans say about the effects of monetary policy on credit risk-taking? *Econometrica,* 82(2), 463-505.

Jokhio, S. H., & Memon, M. A. (2018). Understanding financial ratios and performance indicators: A case of investors in Pakistan. *Journal of Business and Economics Review*, 3(1), 24-35.

Karimi, M., & Rahmani, M. (2017). Financial statement analysis of Iranian companies using ratio analysis. *International Journal of Accounting and Financial Reporting*, 7(1), 305-318.

Khalid, H., Abbas, Z., & Sadiq, M. (2016). Comparative financial statement analysis of selected companies in Pakistan. *Research Journal of Finance and Accounting*, 7(1), 103-110.

Kieso, D. E., Weygandt, J. J., & Warfield, T. D. (2016). *Intermediate accounting (16th ed.). Hoboken, NJ: John Wiley & Sons.*

Kumar, A., & Rastogi, P. (2019). Financial statement analysis using machine learning. In Proceedings of the *10th International Conference on Computational Intelligence and Communication Networks* (pp. 328-333).

Latif, R. A., Isa, M. A. M., Zaharum, Z., & Baharudin, M. A. (2022). Profitability determinants of technology companies: a study from ace market of bursa malaysia. *International Journal of Academic Research in Business and Social Sciences*, 12(6). https://doi.org/10.6007/ijarbss/v12-i6/14034

Lawal, A. I., Abdulsalam, D. A., & Abdulrasheed, A. A. (2017). Exchange rate fluctuations and the financial performance of listed firms in Nigeria. *Research Journal of Finance and Accounting,* 8(17), 26-34.

Lawal, A. I., Oseni, E., Asaleye, A. J., Lawal-Adedoyin, B., & Ojeka-John, R. (2021). Is the stock market efficient? evidence from nonlinear unit root tests for Nigeria. *Asian Economic and Financial Review,* 11(5), 384-395. https://doi.org/10.18488/journal.aefr.2021.115.384.395

Lawal, R. A., & Yusuf, S. A. (2018). The impact of Gross Domestic Product (GDP) on stock market performance in Nigeria. *Journal of Accounting, Finance and Auditing Studies,* 4(2), 85-96.

Liang, Y., Ren, R., & Sun, C. (2020). Earnings quality and stock price crash risk: Evidence from Chinese listed companies. *Journal of Business Research*, 116, 125-135.

Liang, M., Yang, B., Zhang, X., & Zhang, J. (2019). Stock price prediction using machine learning: A review. *International Journal of Financial Studies*, 7(3), 38.

Li, Y., & Li, X. (2020). Using machine learning for customer retention analysis in banking industry. *Journal of Big Data*, 7(1), 56. doi: 10.1186/s40537-020-00330-6

Majeed, M. T., Azam, M., & Amjad, N. (2017). Impact of dividend yield on stock prices: Evidence from Pakistan. *Pakistan Journal of Commerce and Social Sciences, 11(1), 251-270.s*

Muttia, E. E. and Sutrisno, S. (2022). Financial performance and prediction of financial distress in food and beverage companies listed on the indonesia stock exchange. *International Journal of Economics, Business and Management Research, 06(05), 154-167*. https://doi.org/10.51505/ijebmr.2022.6513

National Bureau of Statistics (2021). Consumer Price Index and Inflation

Nsehe, M. (2018, June 7). GTCO, Zenith Bank, And Access Bank Are Nigerian Banks of the Year. Forbes. Retrieved from https://www.forbes.com/sites/mfonobongnsehe/2018/06/07/gtbank-zenith-bank-and-access-bank-are-nigerian-banks-of-the-year/?sh=128f1a9a1b11

Nigeria Stock Exchange. (2021). Market Data. https://www.nse.com.ng/market-data/

Obadiaru, E.D., Obasaju, B.O., Omankhanlen, A.E. and Eyiolorunshe, D.T. (2020) "Dynamic links between the Nigerian equity market and those of selected regional and developed countries," *Heliyon*, 6(9), p. e04782. doi: 10.1016/j.heliyon. 2020.e04782.

Odo, C., Okpara, G. C., Ugwu, S. U., & Onwe, I. O. (2019). Impact of interest rate on stock market performance in Nigeria: A machine learning approach. *Journal of Economics and Sustainable Development*, 10(16), 222-231.

Olayinka, T. O., & Ogundele, O. J. (2018). The impact of company size on financial performance: Evidence from machine learning algorithms. *International Journal of Intelligent Computing and Cybernetics*, 11(2), 213-228.

Olugbenga, A. (2012). Exchange Rate Volatility and Stock Market Behaviour: The Nigerian Experience. https://core.ac.uk/download/pdf/234629265.pdf.

Olugbenga, A. S., & Adetunji, O. I. (2019). Machine learning for stock price prediction in Nigerian stock exchange. *International Journal of Advanced Computer Science and Applications*, 10(2), 157-164.

Ogbebor, P. A., & Okorie, K. C. (2019). An empirical investigation of the capital structure determinants of quoted firms in Nigeria. *African Journal of Accounting, Auditing and Finance*, 8(4), 296-312.

Ogundipe, S. E., & Idowu, A. (2017). Determinants of stock prices in the Nigerian stock exchange. *Research Journal of Finance and Accounting*, 8(19), 2222-5697

Ogunjobi, T., & Ogunnaike, O. O. (2021). Predicting the financial performance of selected companies listed in the Nigeria stock exchange using interest rate data: A machine learning approach. *Data in Brief*, 36, 107023.

Ogundele, O. J., & Olayinka, T. O. (2021). Predicting the future financial performance of companies listed on the Nigeria Stock Exchange using industry trend analysis. *International Journal of Business Forecasting and Marketing Intelligence*, 7(2), 143-160.

Okpala, K. C., & Obi, C. S. (2018). The impact of profitability ratios on stock prices: Evidence from Nigeria. *Journal of Financial Reporting and Accounting*, 16(1), 119-139.

Oladejo, M. A., & Oluwole, O. O. (2020). Firm-specific determinants of profitability of Nigerian listed firms. *Journal of Financial Reporting and Accounting*, 18(4), 543-564.

Oladele, P. O., Adetiloye, K. A., & Oke, M. O. (2020). Market capitalization and financial performance: Evidence from Nigerian listed firms. *Journal of Accounting and Management Information Systems,* 19(2), 311-332.

Olaoye, O. A., & Eriki, P. O. (2019). Market capitalization and financial performance of quoted companies in Nigeria. *Journal of Accounting and Financial Management*, 5(1), 16-26.

Olugbenga, A. O., & Adetayo, A. A. (2018). Effect of interest rate on the financial performance of selected deposit money banks in Nigeria. *Journal of Finance and Investment Analysis*, 7(2), 12-26.

Oyetayo, V. O., Iyiola, O. O., & Oyekunle, R. O. (2015). Dividend policy and financial performance of listed banks in Nigeria. *Journal of Applied Accounting Research,* 16(2), 221-238.

Oyebanjo, O. J., & Oladipupo, A. O. (2019). Exchange rate fluctuations and financial performance of banks in Nigeria. *Journal of Applied Accounting and Taxation*, 6(2), 1-10.

Okafor, G. C., & Obute, G. C. (2018). Neural network analysis of Nigerian stock prices. *International Journal of Advanced Computer Science and Applications*, 9(6), 381-386.

Okoro, O. J., & Akaniro, E. M. (2019). The effect of leverage on the return on equity of listed manufacturing companies in Nigeria. *International Journal of Innovative Research and Advanced Studies*, 6(8), 32-38.

Purwanto, E. and Purwanto, A. D. B. (2020). An investigative study on sustainable competitive advantage of manufacture companies in Indonesia. *Business: Theory and Practice,* 21(2), 633-642. https://doi.org/10.3846/btp.2020.12256

Raza, S. A., Saqib, N., & Masood, A. (2015). Financial analysis of selected companies in Pakistan using vertical analysis technique. *Journal of Finance and Accounting*, 3(1), 1-9.

Shiller, R. J. (1989). Market volatility and macroeconomic fluctuations. *The American Economic Review*, 79(5), 1116-1132.

Sondhi, N., & Kaur, J. (2018). Understanding financial ratios and their significance: An empirical analysis of Indian investors. *International Journal of Applied Business and Economic Research*, 16(3), 311-320.

Talwar, N., Bhatia, V., & Kalia, N. (2020). Financial statement analysis using machine learning: A review. *Journal of Financial Reporting and Accounting*, 18(2), 271-289.

Tan, J., Lo, K., Fang-dao, Q., Zhang, X., & Zhao, H. (2019). Regional economic resilience of resource-based cities and influential factors during economic crises in china. *Growth and Change*, 51(1), 362-381. https://doi.org/10.1111/grow.12352

Wei, B., Yue, J., & Rao, Y. (2017). A deep learning framework for financial time series using stacked auto encoders and long-short term memory. *Plos One*, 12(7), e0180944. https://doi.org/10.1371/journal.pone.0180944

Witten, I. H., Frank, E., Hall, M. A., & Pal, C. J. (2016). Data Mining: Practical Machine Learning Tools and Techniques (3rd ed.). *Morgan Kaufmann Publishers*.

Xia, J., Wu, Z., Zhang, R., Chen, X., & Zhang, R. (2022). Shareholder personal risk and firm risk: an empirical analysis of share pledges and firm debt policies. *Frontiers in Psychology*, 13. https://doi.org/10.3389/fpsyg.2022.1010162

Xu, L. and Cai, F. (2009). Before and after 2000: revenue and high tech valuation. Competitiveness Review: *An International Business Journal*, 19(1), 26-35. https://doi.org/10.1108/10595420910929040

Zairi, M., Ahmed, A. M., & Jabnoun, N. (2005). Trend analysis of financial ratios: A tool for company's performance analysis. Benchmarking: *An International Journal*, 12(5), 443-454.

Zeng, Y., Liu, Z., & Fang, Y. (2020). The impact of Altman Z-score on stock prices: Evidence from Chinese listed companies. *Journal of Risk Research*, 23(1), 116-132.